University of West Bohemia
Faculty of Applied Sciences
Department of Mathematics

BACHELOR THESIS

A Comparison of Different Stochastic Population Models with Regard to Persistence Time

Plzeň 2013 Petra Látalová

Acknowledgement

I would like to thank my supervisor of the bachelor thesis, Ing. Jan Pospíšil Ph.D., for his constant support, suggestions, pieces of advice and guidance. Also I would like to thank my consultant from OU Ostrava, prof. RNDr. Ing. Ivan Křivý, CSc., for his suggestions. My one week internship at OU in Ostrava was supported by the project A-Math-Net - knowledge transfer network in applied mathematics (project no. CZ.1.07/2.4.00/17.0100). This project is co-financed by the European Social Fund and the state budget of the Czech Republic.

Declaration

I declare that I made the bachelor thesis titled A Comparison of Different Stochastic Population Models with Regard to Persistence Time alone and that all materials I have used are cited in the biography.

In Plzeň _____ _____
Signature

Preface

The subject of the bachelor thesis is a comparison of two different stochastic population models with regard to persistence time. We will study the discrete and continuous-time Markov chain models which are commonly used in population biology.

At first we study the theory of stochastic population models where we describe the linear and the general birth and death process. Then the discrete and continuous-time Markov chain models will be discussed. These models take into account the random nature of the individual birth and death process-demographic variability but do not consider random fluctuations of the population in the environment.

Finally we present some comparisons of the two Markov chain models to show that the models agree well.

Plzeň
Czech Republic

Contents

Chapter 1

Introduction

Population models are types of mathematical models that are applied to the study of population dynamics. Models allow a better understanding of how complex interactions and processes work. Modeling of dynamic interactions in nature can provide a manageable way of understanding how numbers change over time or in relation to each other. The models may help to explain the system, to make predictions about behavior and to study the effects of different components.

Population models are used to determine the maximum harvest for agriculturists, to understand the dynamics of biological invasions, and have numerous environmental conservation implications. Population models are also used to understand the spread of parasites, viruses, and disease. The realization of our dependence on environmental health has created a need to understand the dynamic interactions of the Earth's flora and fauna. Methods in population modelling have greatly improved our understanding of ecology and the natural world [7].

Population modelling became of particular interest in the twentieth of the 20th century and is closely connected to the names like biologist Raymond Pearl, physicist Alfred J. Lotka and mathematician Vito Volterra. They formed the nowadays so called deterministic Volterra-Lotka model for competition that applies the logistic equation to two species illustrating their competition. Deterministic population models nowadays more or less follow the ideas of Volterra and Lotka that unfortunately showed to be a weak description of the Earth's processes that are greatly stochastic [8].

Stochastic population models originate to the second half of the 20th century and are nowadays studied rather extensively. Stochastic models are mathematical models used as a tool for estimating probability distributions of potential population outcomes by allowing for random variation in one or more inputs over time. The random variation is usually based on fluctuations observed in historical data for a selected period of time which uses standard time-series techniques. There exist several types of stochastic population models, for example Markov chain or Markov processes models

[4], models described by stochastic differential equations or random dynamical systems etc [5].

In this thesis we focus on discrete or continuous time Markov chain models as they are described by Allen and Allen in their paper [1]. The book Finite Markov Chains and Algorithmic Applications [3] also deals with the theory of Markov chains but is mostly focused on algorithmic applications. In general, there exits many books which deal with the theory of Markov chains but they are usually focused on one specific topic for example as Markov chain Monte Carlo, which is discussed in book Markov Chains: Gibbs field, Monte Carlo Simulations and Queues [2].

In Chapter 2 we focus on theory of stochastic population models. We present preliminaries such as definitions of Markov chains and their basic properties. We consider linear birth and death process and general birth and death process that are widely described in [6].

Chapter 3 is devoted to stochastic birth and death processes. In Section 3.1 we present the discrete-time Markov chain model. In this section we determine the transition matrix, the absorbing state and the mean persistence time. In Section 3.2 we present the continuous-time Markov chain model. We determine the intensity matrix, the absorbing state and the mean persistence time. We will also proof some propositions about the mean persistence time for linear and nonlinear case and also for the liner and nonlinear case of higher order moments.

Chapter 4 is devoted to numerical examples. For pedagogical purposes we present the deterministic solution of the logistic model. The main result of the thesis is a comparison of the discrete-time and the continuous-time model. We show that these two models agree well even for small initial population sizes based on simulations where we present the mean, persistence time and standard deviation for both models. At the end we summarise advantages and disadvantages of the presented models.

In Appendix we present the factorisation method for a 3-diagonal matrix which can be used to proof the solution for the mean persistence time mentioned in Chapter 3 in Section 3.2. We also present the solution of the deterministic equation and we provide some information about the attached MATLAB files.

Chapter 2

Theory of stochastic population models

2.1 Preliminaries

In stochastic models the population size is considered as a random variable dependent on time. Let $J(t)$ denote the random variable for the total population size at time t, N denote the maximum population size. In both discussed cases $J(t)$ is a discrete random variable $J(t) \in \{0, 1, 2, \ldots, N\}$.

Let $\lambda(j)$ denote the birth rate and $\mu(j)$ the death rate of the population when its size is $j \in \{0, 1, \ldots, N\}$. The basic population model expects the birth rate $\lambda(j)$ and the death rate $\mu(j)$ to be continuous and differentiable functions of the population size j. It is also assumed the existence of nonnegative numbers K and N, such that $0 < K < N$ and

$$\lambda(0) = \mu(0) \quad \text{and} \quad \lambda(j) = 0 \quad \text{for} \quad j \geq N, \tag{2.1}$$

$$\lambda(j) > 0 \quad \text{for} \quad j \in (0, N) \quad \text{and} \quad \mu(j) > 0 \quad \text{for} \quad j \in (0, N], \tag{2.2}$$

$$\lambda(j) > \mu(j) \quad \text{for} \quad j \in (0, K), \tag{2.3}$$

$$\lambda(j) < \mu(j) \quad \text{for} \quad j \in (K, N). \tag{2.4}$$

Assumption (2.1) says that when there has no-one been born there cannot die anyone and that there will not be born anyone even if the population size is j when $j \geq N$. Assumption (2.2) shows that there will be some births and deaths when $j \in (0, N)$ and that a death can occure if $j = N$. Assumption (2.3) shows that for $j \in (0, K)$ the population will not extinct. If j does not belong to the interval $(0, K)$ but to the interval $(0, \infty)$ an explosion will occure. Assumption (2.4) says that the population will extinct if $j \in (K, N)$.

In the theory of random processes it is important to mention the following properties and definitions which will be used in the text.

Let (ω, A, P) be a probability space where ω is the state space, A is a $\sigma - algebra$ and P is the probability measure. Let $T \subset \mathbb{R}$ be a set of times. The family of random variables $(X_t, t \in T)$ defined at (ω, A, P) is called the random process. We will consider both discrete-time (e.g.$T \subset \mathbb{N}$) and continuous-time random processes.

Definition 2.1.1. The sequence of integer random variables $\{X_n, n \in \mathbb{N}_0\}$ is called the Markov chain with discrete time and the state space S if

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i) \tag{2.5}$$

for all $n = 0, 1, \dots$ and all $i, j, i_{n-1}, \dots, i_0 \in S$ such that

$$P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) > 0$$

Definition 2.1.2. The system of integer random variables $\{X_t, t \geq 0\}$ defined on the state space $\{\omega, A, P\}$ is called the Markov chain with continuous time and the countable state space $S = \{0, 1, \dots\}$ if

$$P(X_t = j | X_s = i, X_{t_n} = i_n, \dots, X_{t_1} = i_1) = P(X_t = j | X_s = i) \tag{2.6}$$

for all $i, j, i_1, \dots, i_n \in S$ and for all $0 \leq t_1 < t_2 < \dots < t_n < s < t$ for which

$$P(X_s = i, X_{t?n} = i_n, \dots, X_{t_1} = i_1) > 0$$

Markov property means that the probability of a result in the future time $n + 1$ if we know the result in the present time n and results of the past times $n - 1, n - 2, \dots, 0$, is the same as if we know only the result in the present time. The Markov property can be also marked as property without the dependence on the past.

A state j is called accessible from a state i (written $i \to j$) if a system started in state i has a non-zero probability of transitioning into state j at some point. State j is accessible from state i if there exists an integer $n \geqq 0$ such that

$$P(X_n = j | X_{=} i) = p_{ij}^n > 0$$

where p_{ij}^n is the transition probability from state i to state j. For $n = 0$ it means that every state is accessible from itself.

A state i is said to communicate with state j ($i \leftrightarrow j$) if both $i \to j$ and $j \to i$. A set of states C is a communicating class if every pair of states in C communicates with each other, and no state in C communicates with any state not in C. Communication in this sense is an equivalence relation and communicating classes are the equivalence classes of this relation. A communicating class is closed if the probability of leaving the class is zero in other words that if i is in C but j is not, then j is not accessible from i.

A Markov chain called irreducible if it is possible to get to any state from any state.

The time of the first return to the state j $\tau_j(1)$ is a random variable which takes the values $1, 2, \dots$ or value ∞ and marks a random moment where the Markov chain at first gets into the state j after leaving the initial state.

The state j of the Markov chain is called permanent if the chain which comes from j returns back to j with the probability 1 after the finitely many steps i.e.

$$P_j(\tau_j(1) < \infty) = 1.$$

[6]

2.2 Linear birth and death process

Stochastic birth and death processes are described as a population of individuals that are able to duplicate and die. Let $o(h)$ be a residual function such that

$$\lim_{h\to 0} \frac{o(h)}{h} = 0.$$

For simple models we assume that the probability that a new individual is born to an arbitrary individual in the short time interval $(t, t+h]$ is $\lambda h + o(h)$ and that more individuals are born with the probability $o(h)$. In the interval $(t, t+h]$ one individual can die with the probability $\mu h + o(h)$ and more individuals can die with the probability $o(h)$. We also assume that the destinies of the individuals are independent.

The range of the population X_t at time t forms a the continuous-time Markov chain with the following intensities $Q = (q_{i,j})$ [1]

$$\begin{aligned} q_{j,j+1} &= j\lambda, \quad 0 \le j < \infty, \\ q_{j,j-1} &= j\mu, \quad 1 \le j < \infty, \\ q_{jk} &= 0, \quad \text{otherwise}, \\ q_j &= j(\lambda+\mu), \quad 0 \le j < \infty. \end{aligned}$$

The absolute probabilities $p_j(t) = P(X_t = j)$ conforms the system of differential equations

$$\begin{aligned} p_0^{'}(t) &= \mu p_1(t) \\ p_j^{'}(t) &= (j-1)\lambda p_{j-1}(t) - j(\lambda+\mu)p_j(t) + (j+1)\mu p_{j+1}(t), \quad j = 1, 2, \ldots \end{aligned}$$

and we consider the initial condition $p_1(0) = 1$.

To solve this system a generating function method ca be used, see [6] pp. 107-109 to get that for $\lambda = \mu$

$$\begin{aligned} p_0(t) &= \frac{\lambda t}{1+\lambda t} \\ p_j(t) &= \frac{(\lambda t)^{j-1}}{(1+\lambda t)^{j+1}}, \quad 1 \le j < \infty \end{aligned}$$

[1] Absolute intensities q_j appear in the intensity matrix Q on the diagonal as $q_{j,j} = -q_j$.

and for $\lambda \neq \mu$
where for simplifying we use

$$A(t) = \frac{1 - e^{(\lambda-\mu)t}}{\mu - \lambda e^{(\lambda-\mu)t}}$$

and

$$\begin{aligned} p_0(t) &= \mu A(t) \\ p_j(t) &= (1 - \lambda A(t))(1 - \mu A(t))(\lambda A(t))^{j-1}, \quad j \geq 1. \end{aligned}$$

The matrix of transition probabilities Q^* among the states of the nested discrete-time Markov chain shows that the zero state is an absorbing state. This means that the population that has once extinct cannot recover any more.

$$\mathbf{Q}^* = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ \frac{\mu}{(\lambda+\mu)} & 0 & \frac{\lambda}{(\lambda+\mu)} & 0 & \dots \\ 0 & \frac{\mu}{(\lambda+\mu)} & 0 & \frac{\lambda}{(\lambda+\mu)} & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

The probability of population extinction is therefore

$$P(X_t = 0) = p_0(t) = \begin{cases} \frac{\lambda t}{(1+\lambda t)}, & \text{if } \lambda = \mu, \\ \frac{1-e^{(\lambda-\mu)t}}{(\mu-\lambda e^{(\lambda-\mu)t}}\mu, & \text{if } \lambda \neq \mu. \end{cases}$$

Using the limit transition for $t \to \infty$ we find out that the population extinction is

$$\lim_{t\to\infty} p_0(t) = \begin{cases} 1, & \text{if } \lambda \leq \mu, \\ \frac{\mu}{\lambda}, & \text{if } \lambda > \mu. \end{cases}$$

[6]

2.3 General birth and death process

In the general birth and death process as well as in the linear birth and death process we assume the population of individuals to be able to duplicate and die. The population range at time t forms a continuous-time Markov chain with the following intensities:

$$\begin{aligned} q_{j,j+1} &= \lambda_j, \quad j = 0, 1, \dots \\ q_{j,j-1} &= \mu_j, \quad j = 1, 2, \dots \\ q_{jk} &= 0 \quad \text{otherwise} \\ q_0 &= \lambda_0 \\ q_j &= \lambda_j + \mu_j, \quad j = 1, 2, \dots . \end{aligned}$$

The intensity matrix is

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \dots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \dots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

For better illustration we give a detail of the matrix Q in the neighbourhood of the j-th row

$$\mathbf{Q} = \begin{pmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \dots \\ \ddots & \mu_{j-1} & -(\lambda_{j-1}+\mu_{j-1}) & \lambda_{j-1} & 0 & 0 & \dots \\ \ddots & 0 & \mu_j & -(\lambda_j+\mu_j) & \lambda_j & 0 & \dots \\ \ddots & 0 & 0 & \mu_{j+1} & -(\lambda_{j+1}+\mu_{j+1}) & \lambda_{j+1} & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

Our aim is to find absolute probabilities $p_j(t), j \in S$, with the fixed initial state i which means for the initial distribution $p_i = 1, p_j = 0, j \neq i$. The system of differential equations for absolute probabilities $p_j(t)$ we derive under the same conditions as the corresponding system of prospective Kolmogorov differential equations ([6] pp. 82-83).

$$\begin{aligned} p_0^{'}(t) &= -\lambda_0 p_0(t) + \mu_1 p_1(t) \\ p_j^{'}(t) &= \lambda_{j-1} p_{j-1}(t) - (\lambda_j + \mu_j) p_j(t) + \mu_{j+1} p_{j+1}(t), \quad j = 1, 2, ..., \end{aligned}$$

with initial condition $p_i(0) = 1, p_j = (0) = 0, j \neq i$.

Considering these initial conditions the absolute probabilities match the i-th row of matrix $P(t)$.

Furthermore, we restrict ourselves to calculating the limit probabilities for which we use the following theorem.

Theorem 2.3.1. *Let $\{X_t, t \geq 0\}$ be a continuous-time Markov chain with the intensity matrix Q and the nested chain $\{Y_n, n \in \mathbb{N}_0\}$, which is irreducible with all states permanent. Let $\pi = \{\pi_j, j \in S\}$ where $\pi_j > 0$ for all $j \in S$, $\sum_{j \in S} \pi_j = 1$, is the nontrivial solution of the system equations*

$$\boldsymbol{\pi}^T \boldsymbol{Q} = \mathbf{0}^T.$$

Then for the transition probabilities and the absolute probabilities in the chain $\{X_t, t \geq 0\}$ are

$$\lim_{t \to \infty} p_{i,j}(t) = \pi_j \quad \textit{for all} \quad i, j \in S$$

$$\lim_{t \to \infty} p_j(t) = \pi_j \quad \textit{for all} \quad j \in S.$$

Proof. See [6] pp. 94. □

We will assume that all intensities $\lambda_j, j \geq 0$, and $\mu_j, j > 0$ are positive. Then the transition probability matrix of a nested chain is

$$\mathbf{Q}^* = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ \frac{\mu_1}{\lambda_1+\mu_1} & 0 & \frac{\lambda_1}{\lambda_1+\mu_1} & 0 & \dots \\ 0 & \frac{\mu_2}{\lambda_2+\mu_2} & 0 & \frac{\lambda_2}{\lambda_2+\mu_2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

It can be seen that the nested chain is irreducible. Whether all the states of the nested chain are permanent can be decided using Theorem (2.3.1) or we may also find out if the stationary distribution exists.

Theorem 2.3.2. *In an irreducible chain with the state space $S = 0, 1, ...$, all states are permanent if and only if the only solution of the equations*

$$x_i = \sum_{j=1}^{+\infty} q_{ij}^* x_j, i = 1, 2, ... \tag{2.7}$$

on the iterval $[0, 1]$ is a trivial solution $x_i = 0, i = 1, 2, ...$ All states are transitional if and only if (2.7) has in $[0, 1]$ a nontrivial solution.

Proof. See [6] pp. 49. □

Theorem 2.3.3. *Let $\{X_t, t \geq 0\}$ be a continuous time Markov chain with the intensity matrix Q and a nested chain $\{Y_n, n \in \mathbb{N}_0\}$ which is irreducible with all states permanent. Then there exists an invariant measure η of the process $\{X_t, t \geq 0\}$ which is determined uniquely (excluding a multiplicative constant) as the nontrivial solution of the system*

$$\boldsymbol{\eta}^T \boldsymbol{Q} = \mathbf{0}^T.$$

Moreover

$$\pi_j = \frac{\eta_j}{\sum\limits_{j \in S} \eta_j}$$

is the stational distribution of the process $\{X_t, t \geq 0\}$.

Proof. See [6] pp. 92-93. □

In our case it is sufficient to show that $\sum\limits_{j \in S} \pi_j q_j < \infty$ where $\{\pi_j\}$ is the non-trivial solution of the equation $\boldsymbol{\pi}^T \mathbf{Q} = \mathbf{0}^T$ because the proof of Theorem (2.3.3) (See [6] pp. 92-93) implies that $\{\pi_j q_j\}$ is an invariant measure in the nested chain.

Being the conditions of our Theorem (2.3.1) fulfilled we can find the limit probabilities $\pi_j = \lim_{n\to\infty} p_j(t)$ as a solution of the system $\boldsymbol{\pi}^T\mathbf{Q} = \mathbf{0}^T$ which satisfies the conditions $\pi_j > 0, \sum_{j\in S} \pi_j = 1$.

The considered system is

$$\begin{aligned} -\lambda_0\pi_0 + \mu_1\pi_1 &= 0 \\ -\lambda_{j-1}\pi_{j-1} - (\lambda_j + \mu_j)\pi_j + \mu_{j+1}\pi_{j+1} &= 0, \quad \geq 1. \end{aligned}$$

If we put

$$\mu_j\pi_j - \lambda_{j-1}\pi_{j-1} = K_j, \quad \geq 1$$

we can rewrite this system in the following form

$$\begin{aligned} K_1 &= 0 \\ K_{j+1} - K_j &= 0, \quad j \geq 1. \end{aligned}$$

Apparently $K_j = 0$ for $j \geq 1$, so

$$\pi_j = \frac{\lambda_{j-1}}{\mu_j}\pi_{j-1} = \frac{\lambda_{j-1}\lambda_{j-2}}{\mu_j\mu_{j-1}}\pi_{j-2} = ... = \frac{\lambda_{j-1}\lambda_{j-2}...\lambda_0}{\mu_j\mu_{j-1}...\mu_1}\pi_0 = \rho_j\pi_0,$$

where we have denoted

$$\rho_j = \frac{\lambda_{j-1}\lambda_{j-2}...\lambda_0}{\mu_j\mu_{j-1}...\mu_1}\pi_0, \quad j \geq 1.$$

If we define $\rho_0 = 1$, we can write

$$\pi_j = \rho_j\pi_0, \quad j \geq 0.$$

The solution $\{\pi_j, j \geq 0\}$ will be the probability distribution if and only if $\sum_{k=1}^{\infty} \rho_k < \infty$. Hence

$$\pi_j = \rho_j \left(\sum_{k=0}^{\infty} \rho_k \right)^{-1} \text{ and } \pi_j > 0, j \geq 0.$$

If $\lambda_0 = 0$ then 0 will be an absorbing state and the nested chain will be reducible (will be $q^*_{00=1}$). The probability that the population which has just i individuals at the beginning will ever extinct is the same as the probability that the nested chain which was in the state i at the beginning will ever end in the absorbing state 0. The probability of extinction that we were looking for is $\lim_{n\to\infty} p_{i0}(t)$. If we denote it u_i we can determine using the following theorem.

The variable T is a set of transitional states, τ is a random variable $\tau = inf\{n > 0 : X_n \notin T\}$ which marks the time output from the set of transitional states T, T^c is the closed set of permanent states and X_τ is the state in which the chain comes as soon as it leaves the set of transitional states.

Theorem 2.3.4. *For probability u_{ij} defined as $u_{ij} = P_i(X_\tau = j), \quad i \in T, j \in T^c$ we have that*

$$u_{ij} = p_{ij} + \sum_{v \in T} p_{iv} u_{vj}, \quad i \in T, j \in T^c.$$

Proof. See [6] pp. 43-44. □

Let $J_0 = 0$ and let $J_1, J_2, \ldots$ be consecutive time moments where occur transitions between the states of the Markov chain i.e.

$$\begin{aligned} J_1 &= inf\{t > 0 : X_t \neq X_0\}, \\ J_2 &= inf\{t > J_1 : X_t \neq X_{J_1}\}, \\ &\ldots \\ J_{n+1} &= inf\{t > J_n : X_t \neq X_{J_n}\}, n \geq 0. \end{aligned}$$

The times between individual transitions are $S_1 = J_1, S_2 = J_2 - J_1$ (if $J_1 < \infty$, else $S_2 = \infty$) and so on. Apparently

$$J_n = \sum_{k=1}^{\infty} S_k \quad n \geq 1.$$

Then we can denote

$$\xi = sup J_n = \sum_{k=1}^{\infty} S_k.$$

Definition 2.3.1. Homogenous Markov chain with permanent states is called regular if

$$Pi(\xi = \infty) = 1 \; \forall i \in S.$$

The random variable ξ is called time to explosion.

If a process is regular it means that for any finite interval $(0, t)$ there will be with probability one only a finite number of transitions among the states of the chain. If there was an infinite number of transitions in the finite interval, the process would explode.

The probability of extinction can be determined according to the Theorem (2.3.4). with probability q_{ij}^* as the solution of the system

$$u_i = q^*_{i0} + \sum_{k=0}^{\infty} q^*_{ik} u_k, \quad i = 1, 2, ...,$$

or as the solution of the system

$$u_1 = \frac{\mu_1}{\lambda_1 + \mu_1} + \frac{\lambda_1}{\lambda_1 + \mu_1} u_2$$
$$u_i = \frac{\mu_i}{\lambda_i + \mu_i} u_{i-1} + \frac{\lambda_i}{\lambda_i + \mu_i} u_{i+1}, \quad i \geq 2.$$

It is also worth mentioning that the sufficient condition for regularity of the chain is

$$\sum_{k=0}^{\infty} \frac{1}{\lambda_j} = \infty.$$

This condition ensures regularity in the growth process in which population size only increases. But now it is no longer a necessary condition. [6]

Chapter 3

Stochastic birth and death processes

3.1 Discrete-time Markov chain model

The discrete-time Markov chain model supposes both time and population size to be discrete-valued. Let v be a fixed time interval and $t \in \{0, \triangle h, \triangle 2h\}$. We also assume $\triangle h$ to be sufficiently small so that there occurs at most one change during the time interval $\triangle h$, either a birth or a death.

When we have the population size j probability that there occurs a birth is $\lambda_j \triangle h$ and probability that there occures a death is $\mu_j \triangle h$. For the discrete set of points $j \in \{0, 1, 2, \ldots, N\}$ we expect $\lambda_j \triangle h$ and $\mu_j \triangle h$ to satisfy assumption (2.1),(2.2),(2.3) and (2.4).

Let the probabilities associated with $J(t)$ be denoted as

$$p_j = Prob\{J(t) = j\}, j = 0, 1, \ldots, N,$$

and

$$p(t) = (p_0(t), p1(t), \ldots, p_N(t))^T.$$

The transition probabilities we denote as

$$p_{ij}(\triangle h) = Prob\{J(t + \triangle h) = j | J(t) = i\},$$

where

$$\begin{aligned} p_{j,j-1}(\triangle h) &= \mu_j \triangle h \quad \text{for} \quad j \in \{1, 2, \ldots, N\} \\ p_{j,j+1}(\triangle h) &= \lambda_j \triangle h \quad \text{for} \quad j \in \{0, 1, 2, \ldots, N-1\} \\ p_j \triangle h &= 1 - [\lambda_j + \mu_j] \triangle h \quad \text{for} \quad j \in \{0, 1, 2, \ldots, N\} \\ p_{j,k}(\triangle h) &= 0 \quad \text{otherwise.} \end{aligned}$$

Then $P_j(t+\triangle h)$ satisfies

$$\begin{aligned}
p_0(t+\triangle h) &= p_0(t) + \mu_1 \triangle h p_1(t) \quad \text{for} \quad j=0, \\
p_j(t+\triangle h) &= \lambda_{j-1}\triangle h p_{j-1}(t) + (1-[\lambda_j+\mu_j]\triangle h)p_j(t) + \mu_{j+1}\triangle h p_{j+1}(t) \\
&\quad \text{for} \quad j=1,2,\ldots,N-1, \\
p_N(t+\triangle h) &= \lambda_{N-1}\triangle h p_{N-1}(t) + (1-\mu_N\triangle h)p_N(t) \quad \text{for} \quad j=N.
\end{aligned}$$

The difference equations project forward in time. They can be also expressed in a matrix form as

$$p(t+\triangle h) = Pp(t), p_{j_0}(0) = 1, \tag{3.1}$$

where the matrix P, $P = (p_{ij}(\triangle h))$ is the transition matrix:

$$
\mathbf{P} = \begin{pmatrix}
1 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots \\
\mu_1 \triangle h & 1 - [\lambda_1 + \mu_1] \triangle h & \lambda_1 \triangle h & \dots & \dots & \dots & \dots & \dots & \dots \\
0 & \mu_2 \triangle h & 1 - [\lambda_2 + \mu_2] \triangle h & \ddots & \dots & \dots & \dots & \dots & \dots \\
0 & 0 & \mu_3 \triangle h & \ddots & \lambda_{j-1} \triangle h & \dots & \dots & \dots & \dots \\
\dots & \dots & \dots & \ddots & 1 - [\lambda_j + \mu_j] \triangle h & \ddots & \dots & \dots & \dots \\
\dots & \dots & \dots & \dots & \mu_{j+1} \triangle h & \ddots & \lambda_{N-3} \triangle h & 0 & 0 \\
\dots & \dots & \dots & \dots & \dots & \ddots & 1 - [\lambda_{N-2} + \mu_{N-2}] \triangle & \lambda_{N-2} \triangle h & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \mu_{N-1} \triangle h & 1 - [\lambda_{N-1} + \mu_{N-1}] \triangle h & \lambda_{N-1} \triangle h \\
\dots & \dots & \dots & \dots & \dots & \dots & 0 & \mu_N \triangle h & 1 - \mu_N \triangle h
\end{pmatrix}
$$

For our calculations we need to ensure that the matrix P is a stochastic matrix, which means a nonnegative matrix where the row elements sum to one. Because of this we assume that

$$max_{j\in\{1,2,\ldots,N\}}\{[\lambda_j+\mu_j]\triangle h\}\leqslant 1.$$

Zero si the only absorbing state in the discrete-time model, $p_{00}(\triangle h)=1$ which is equivalent for

$$\lim_{t\to\infty} p_0(t)=1$$

or we can say that the extinction of the population occurs with probability one.

Let T be the random variable for the time until the population extinction. The distribution of T depends on the initial population size so we will denote this dependence by T_{y_0}. Let t_{y_0} denote the expected time until extinction or the mean persistence time for the population size y_0. This can be expressed as

$$\tau_{y_0}=E(T_{y_0}).$$

Let $\tau^r_{y_0}=E(T^r_{y_0})$, $r>1$ denote the r-th moment. Because $\tau_0=0$ and $\tau^r_0=0$ for $r>1$ zero is an absorbing state. For an initial population size $y_0\in\{1,2,\ldots,N\}$ let the moment vectors for the persistence time denote as $\tau=\{\tau_1,\ldots,\tau_N\}$ and $\tau^r=\{\ldots tau^r_1,\ldots,\tau^r_N\}$ for $r>1$.

For a small period of time $\triangle h$ in the discrete time model $p(t+\triangle h)=Pp(t)$, $p_{j_0}(0)=1$, there can occurs either a birth or a death or no change in the population size. The mean persistence time satisfies these difference equations

$$\tau_j=\lambda_j\triangle h(\tau_{j+1}+\triangle h)+(1-[\lambda_j+\mu_j]\triangle h)(\tau_j+\triangle j)+\mu_j\triangle h(\tau_{j-1}+\triangle h) \tag{3.2}$$

where $j=1,2,\ldots,N$. After multiplying the brackets

$$\tau_j=\lambda_j\triangle h\tau_{j+1}+\lambda_j\triangle h^2+\tau_j+\triangle h-\lambda_j\triangle h\tau_j-\lambda_j\triangle h^2$$

$$-\mu_j\triangle h\tau_j-\mu_j\triangle h^2+\mu_j\triangle h\tau_{j-1}+\mu_j\triangle h^2$$

and substracting $\tau_j,\lambda_j\triangle h^2$ and $\mu_j\triangle h^2$, we get

$$0=\lambda_j\triangle h\tau_{j+1}+\triangle h-\lambda_j\triangle h\tau_j-\mu_j\triangle h\tau_j+\mu_j\triangle h\tau_{j-1}$$

where after dividing the whole equation by $\triangle h$

$$-1=\lambda_j\tau_{j+1}-\lambda_j\tau_j-\mu_j\tau_j+\mu_j\tau_{j-1}$$

we obtain the following simplified form of the difference equations.

$$\lambda_j\tau_{y+1}-[\lambda_j+\mu_j]\tau_y+\mu_j\tau_{y-1}=-1,\quad j=1,\ldots,N. \tag{3.3}$$

This simplified relationship is based on the backward difference equation as opposed to the forward difference equations mentioned in (3.3).

[1]

3.2 Continuous-time Markov chain model

In the continuous-time Markov chain model $t \in [0, \infty)$ and $J(t) \in \{0, 1, 2, \ldots, N\}$. For sufficiently small $\triangle h$ the infinitesimal or endlessly small transition probabilities $p_{ij}(\triangle t)$ are similar to those in the discrete-time Markov chain model. Also the transition matrix for this model assume

$$\begin{aligned}
p_{j,j-1}(\triangle h) &= \mu_j \triangle h + o(h) \quad \text{for} \quad j \in \{1, 2, \ldots, N\} \\
p_{j,j+1}(\triangle h) &= \lambda_j \triangle h + o(h) \quad \text{for} \quad j \in \{0, 1, 2, \ldots, N-1\} \\
p_j \triangle h &= 1 - [\lambda_j + \mu_j] \triangle h + o(h) \quad \text{for} \quad j \in \{0, 1, 2, \ldots, N\} \\
p_{j,k}(\triangle h) &= o(h) \quad \text{otherwise.}
\end{aligned}$$

It can be shown that probabilities

$$p_y(t) = Prob\{J(t) = y\}.$$

as $\triangle h \to 0$ satisfies the forward Kolmogorov difference equations

$$\begin{aligned}
\frac{dp_y(t)}{dt} &= \lambda_{j-1} p_{y-1}(t) - [\lambda_j + \mu_j] p_y + \mu_{j+1} p_{j+1}(t), \quad y \in 1, 2, \ldots, N, \\
\frac{dp_0(t)}{dt} &= \lambda_1 p_1(t).
\end{aligned}$$

We can rewrite this system in the following matrix form as

$$\frac{dp(t)}{dt} = Qp, p_{j_0} = 1. \tag{3.4}$$

Matrix $Q = (q_{ij})$ is here the infinitesimal generator matrix:

$$
Q = \begin{pmatrix}
0 & \mu_1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & -[\lambda_1+\mu_1] & \mu_2 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \lambda_1 & -[\lambda_2+\mu_2] & \mu_3 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \ddots & \ddots & \ddots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \lambda_{j-1} & -[\lambda_j+\mu_j] & \mu_{j+1} & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \ddots & \ddots & \ddots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \lambda_{N-3} & -[\lambda_{N-2}+\mu_{N-2}] & \mu_{N-1} & \cdots & \cdots & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_{N-2} & -[\lambda_{N-1}+\mu_{N-1}] & \mu_N & \cdots & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{N-1} & -\mu_N & \cdots & \cdots
\end{pmatrix}
$$

Remark 3.2.1. By substracting $Ip(t)$ from both sides of (3.1) and after dividing through by $\triangle h$, where the limit would be taken as $\triangle h \to 0$ then we get

$$\lim_{\triangle h \to 0} \frac{P - I}{\triangle h} = Q.$$

The system of difference equations (3.1) may be considered a discrete-time approximation to the difference system (3.4). Approximation (3.1) have similar results as the continuous-time model(3.4) if $\triangle h$ is sufficiently small. Zero is the only absorbing state in (3.4) because $q_{00} = 0$. We can also see that

$$\lim_{t\to\infty} p_0(t) = 1.$$

Continuous-time model (3.4) satisfies the relationship

$$\tau_j = \lambda_j \triangle h(\tau_{j+1} + \triangle h) + (1 - [\lambda_j + \mu_j]\triangle h)(\tau_j + \triangle h) + \mu_j \triangle h(\tau_{j-1} + \triangle h) + o(\triangle h)$$

which is very similar to relationship in (3.2) . For $o(\triangle h) \to 0$ the only difference is the addition of a term of $o(\triangle h)$ to the right side of the equation.[1] The mean persistence time therefore satisfies the same relationship as for the discrete time i.e. equation (3.3). This equation follows from the backward Kolmogorov differential equations.

The equation (3.3) can be rewritten as a single matrix equation

$$D\tau = -\mathbf{1}$$

where $\mathbf{1} = (1, \ldots, 1)^T, \tau = (\tau_1, \ldots, \tau_N)$ and

$$\mathbf{D} = \begin{pmatrix} -[\lambda_1 + \mu_1] & \lambda_1 & 0 & \ldots & 0 & 0 & 0 \\ \mu_2 & -[\lambda_2 + \mu_2] & \lambda_2 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \mu_{N-1} & -[\lambda_{N-1} + \mu_{N-1}] & \lambda_{N-1} \\ 0 & 0 & 0 & \ldots & 0 & \mu_N & -\mu_N \end{pmatrix} \tag{3.5}$$

Matrix D is an irreducible diagonally dominant matrix and therefore nonsingular. The solution for the mean persistence time is

$$\tau = -D^{-1}\mathbf{1} \tag{3.6}$$

which can be prooved using the factorisation method described in Appendix A.1. Because the matrix D is tri-diagonal we can get an explicit solution for τ. If we suppose $J(0) = j$ for $j \in \{1, 2, \ldots, N\}$, then the mean persistence time satisfies [1]

$$\tau_j = \begin{cases} \frac{1}{\mu_1} + \sum_{i=2}^{N} \frac{\lambda_1 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i}, & \text{for } j = 1, \\ \tau_1 + \sum_{s=1}^{j-1} \left[\frac{\mu_1 \ldots \mu_s}{\lambda_1 \ldots \lambda_s} \sum_{i=s+1}^{N} \frac{\lambda_1 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i} \right], & \text{for } j = 2, \ldots, N. \end{cases} \tag{3.7}$$

The equation above can be also simplified in the following form

$$\tau_j = \begin{cases} \frac{1}{\mu_1} + \sum_{i=2}^{N} \frac{\lambda_1 \ldots \lambda_{i-1}}{\mu_1 \ldots \mu_i}, & \text{for } j = 1, \\ \tau_1 + \sum_{s=1}^{j-1} \left[\frac{1}{\lambda_s} \sum_{i=s+1}^{N} \frac{\lambda_s \ldots \lambda_{i-1}}{\mu_{s+1} \ldots \mu_i} \right], & \text{for } j = 2, \ldots, N. \end{cases} \tag{3.8}$$

[1] In fact it is added three times.

Example 3.2.1. In a linear case where

$$\lambda_j = j\lambda$$
$$\mu_j = j\mu$$

the formula for the mean persistence time satisfies

$$\tau_j = \begin{cases} \frac{1}{\mu} + \sum_{i=2}^{N} \frac{(i-1)!\lambda^{i-1}}{i!\mu^i}, & \text{for } j = 1, \\ \tau_1 + \sum_{s=1}^{j-1} \left[\frac{s!\mu^s}{s!\lambda^s} \sum_{i=s+1}^{N} \frac{(i-1)!\lambda^{i-1}}{i!\mu^i} \right], & \text{for } j = 2, \dots, N. \end{cases}$$

which can be also simplified in this form

$$\tau_j = \begin{cases} \frac{1}{\mu} + \sum_{i=2}^{N} \frac{(i-1)!\lambda^{i-1}}{i!\mu^i}, & \text{for } j = 1, \\ \tau_1 + \sum_{s=1}^{j-1} \left[\frac{1}{\lambda^s} \sum_{i=s+1}^{N} \frac{\lambda^{i-1}}{i\mu^{i-s}} \right], & \text{for } j = 2, \dots, N. \end{cases}$$

Proposition 3.2.1. *Let D and τ be given by (3.5) and (3.8). Then $D\tau = -\mathbf{1}$.*

Proof. We will multiply the matrix D (3.5) with τ for which we use the simplified formula (3.8).
Let us denote

$$\gamma_1 = \frac{1}{\lambda_j} \sum_{i=2}^{N} \frac{\lambda_1 \dots \lambda_{i-1}}{\mu_2 \dots \mu_i}$$

then from the formula (3.8) we can express τ_1 as

$$\tau_1 = \frac{1}{\mu_1} + \frac{\lambda_1}{\mu_1}\gamma_1.$$

From the formula (3.8) we also know that for $j = 2, \dots, N-1$

$$\begin{aligned} \tau_j &= \tau_1 + \sum_{s=1}^{j-1} \left[\frac{1}{\lambda_s} \sum_{i=s+1}^{N} \frac{\lambda_s \dots \lambda_{i-1}}{\mu_{s+1} \dots \mu_i} \right] \\ \tau_{j+1} &= \tau_1 + \sum_{s=1}^{j} \left[\frac{1}{\lambda_s} \sum_{i=s+1}^{N} \frac{\lambda_s \dots \lambda_{i-1}}{\mu_{s+1} \dots \mu_i} \right] = \tau_j + \sum_{s=j}^{j} \left[\frac{1}{\lambda_s} \sum_{i=s+1}^{N} \frac{\lambda_s \dots \lambda_{i-1}}{\mu_{s+1} \dots \mu_i} \right] \\ &= \tau_j + \frac{1}{\lambda_j} \sum_{i=j+1}^{N} \frac{\lambda_j \dots \lambda_{i-1}}{\mu_{j+1} \dots \mu_i} \end{aligned}$$

Let us denote

$$\gamma_j = \frac{1}{\lambda_j} \sum_{i=j+1}^{N} \frac{\lambda_j \dots \lambda_{i-1}}{\mu_{j+1} \dots \mu_i}$$

for $j = 1, \ldots, N-1$.
Then we can write

$$\tau_{j+1} = \tau_j + \gamma_j$$

where $j = 1, \ldots, N-1$ and τ is given by (3.8).
We start with the verification for index 1:

$$\begin{aligned}
(D\tau)_1 &= (-\lambda_1 - \mu_1)\tau_1 + \lambda_1\tau_2 \\
&= (-\lambda_1 - \mu_1)\tau_1 + \lambda_1(\tau_1 + \gamma_1) \\
&= -\mu_1\tau_1 + \lambda_1\gamma_1 \\
&= -\mu_1\left(\frac{1}{\mu_1} + \frac{\lambda_1}{\mu_1}\gamma_1\right) + \lambda_1\frac{1}{\lambda_1}\sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_2\ldots\mu_i} \\
&= -1 - \lambda_1\frac{1}{\lambda_1}\sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_2\ldots\mu_i} + \sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_2\ldots\mu_i} \\
&= -1 - \sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_2\ldots\mu_i} + \sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_2\ldots\mu_i} \\
&= -1.
\end{aligned}$$

Then we continue with the $j-th$ index for $j = 2, \ldots, N-1$:

$$\begin{aligned}
(D\tau)_j &= \mu_j\tau_{j-1} - (\lambda_j + \mu_j)\tau_j + \lambda_j\tau_{j+1} \\
&= -\mu_j(\tau_j - \tau_{j-1}) + \lambda_j(\tau_{j+1} - \tau_j) \\
&= -\mu_j\gamma_{j-1} + \lambda_j\gamma_j \\
&= -\mu_j\frac{1}{\lambda_{j-1}}\sum_{i=j}^{N}\frac{\lambda_{j-1}\ldots\lambda_{i-1}}{\mu_j\ldots\mu_i} + \lambda_j\frac{1}{\lambda_j}\sum_{i=j+1}^{N}\frac{\lambda_j\ldots\lambda_{i-1}}{\mu_{j+1}\ldots\mu_i} \\
&= -\mu_j\frac{1}{\lambda_{j-1}}\left(\frac{\lambda_{j-1}}{\mu_j} + \sum_{i=j+1}^{N}\frac{\lambda_{j-1}\ldots\lambda_{i-1}}{\mu_j\ldots\mu_i}\right) + \sum_{i=j+1}^{N}\frac{\lambda_j\ldots\lambda_{i-1}}{\mu_{j+1}\ldots\mu_i} \\
&= -1 - \mu_j\frac{1}{\lambda_{j-1}}\frac{\lambda_{j-1}}{\mu_j}\sum_{i=j+1}^{N}\frac{\lambda_j\ldots\lambda_{i-1}}{\mu_{j+1}\ldots\mu_i} + \sum_{i=j+1}^{N}\frac{\lambda_j\ldots\lambda_{i-1}}{\mu_{j+1}\ldots\mu_i} \\
&= -1.
\end{aligned}$$

And finally we show the verification for the $N-th$ index:

$$\begin{aligned}(D\tau)_N &= \mu_N\tau_{N-1} - \mu_N\tau_N \\ &= -\mu_N(\tau_N - \tau_{N-1}) \\ &= -\mu_N\gamma_{N-1} \\ &= -\mu_N\frac{1}{\lambda_{N-1}}\sum_{i=N}^{N}\frac{\lambda_{N-1}\ldots\lambda_{i-1}}{\mu_N\ldots\mu_i} \\ &= -\mu_N\frac{1}{\lambda_{N-1}}\frac{\lambda_{N-1}}{\mu_N} \\ &= -1\end{aligned}$$

□

From the continuous-time model we can derive an expression for τ_j^r. The rth-order moment can be expressed in terms of the $(r-1)$th-order moment

$$\lambda_j\tau_{j+1}^r + [\lambda_j + \mu_j]\tau_j^r + \mu_j\tau_{j-1}^r = -r\tau_j^r \tag{3.9}$$

for $j = 1, 2, \ldots, N$ (Goel and Richter-Dyn, 1974; Nisbet and Gurney, 1892; Norden, 1982).

This relationship (3.9) follows from the backward Kolmogorov differential equations. If $r = 1$ the equation (3.9) can be reduced to the equation derived from the mean persistence time in (3.3) where $\tau_y^1 = \tau_y$. The difference equation (3.9) can be also expressed in the matrix form as $D\tau^r = -r\tau^{r-1}$ where the solution is

$$\tau^r = -rD^-1\tau^{r-1}. \tag{3.10}$$

Recursively we can find the r-th order moments by applying the operator $-rD^{-1}$ to the $(r-1)$ moment. Following these operations we can derive an explicit solution for ${\tau_j}^r$ as in [1]

$$\tau_j^r = \begin{cases} \frac{r\tau_1^{r-1}}{\mu_1} + r\sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i}\tau_i^{r-1}, & \text{for } j = 1, \\ \tau_1^r + r\sum_{s=1}^{j-1}\left[\frac{\mu_1\ldots\mu_s}{\lambda_1\ldots\lambda_s}\sum_{i=s+1}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i}\tau_i^{r-1}\right], & \text{for } j = 2, \ldots, N. \end{cases} \tag{3.11}$$

The simplified form of this formula is

$$\tau_j^r = \begin{cases} \frac{r\tau_1^{r-1}}{\mu_1} + r\sum_{i=2}^{N}\frac{\lambda_1\ldots\lambda_{i-1}}{\mu_1\ldots\mu_i}\tau_i^{r-1}, & \text{for } j = 1, \\ \tau_1^r + r\sum_{s=1}^{j-1}\left[\frac{1}{\lambda_s}\sum_{i=s+1}^{N}\frac{\lambda_s\ldots\lambda_{i-1}}{\mu_{s+1}\ldots\mu_i}\tau_i^{r-1}\right], & \text{for } j = 2, \ldots, N. \end{cases} \tag{3.12}$$

Proposition 3.2.2. *Let D and τ^r be given by (3.5) and (3.12). Then $D\tau^r = -r\tau^{r-1}$.*

Proof. Because the previous Proposition was just a special case of this one the Proof will be very similar. We will multiply the matrix D (3.5) with τ^r for which we use the simplified formula (3.12).
Let us denote

$$\gamma_1 = r\frac{1}{\lambda_j}\sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_2\dots\mu_i}$$

then from the formula (3.8) we τ_1^r as

$$\tau_1^r = \frac{r\tau_1^r}{\mu_1} + \frac{\lambda_1}{\mu_1}\gamma_1.$$

From the formula (3.12) we know that

$$\begin{aligned}
\tau_j^r &= \tau_1^r + r\sum_{s=1}^{j-1}\left[\frac{1}{\lambda_s}\sum_{i=s+1}^{N}\frac{\lambda_s\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_{s+1}\dots\mu_i}\right]\\
\tau_{j+1}^r &= \tau_1^r + r\sum_{s=1}^{j}\left[\frac{1}{\lambda_s}\sum_{i=s+1}^{N}\frac{\lambda_s\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_{s+1}\dots\mu_i}\right] = \tau_i^{r-1} + r\sum_{s=j}^{j}\left[\frac{1}{\lambda_s}\sum_{i=s+1}^{N}\frac{\lambda_s\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_{s+1}\dots\mu_i}\right]\\
&= \tau_j^r + r\frac{1}{\lambda_j}\sum_{i=j+1}^{N}\frac{\lambda_j\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_{j+1}\dots\mu_i}
\end{aligned}$$

Let us denote

$$\gamma_j = r\frac{1}{\lambda_j}\sum_{i=j+1}^{N}\frac{\lambda_j\dots\lambda_{i-1}\tau_i^{r-1}}{\mu_{j+1}\dots\mu_i}$$

for $j = 1,\dots,N-1$.
Then we can write

$$\tau_{j+1}^{r-1} = \tau_j^r + \gamma_j,$$

where $j = 1,\dots,N-1$.
We start with the verification for index 1:

$$\begin{aligned}
(D\tau^r)_1 &= (-\lambda_1-\mu_1)\tau_1^r + \lambda_1\tau_2^r\\
&= (-\lambda_1-\mu_1)\tau_1^r + \lambda_1(\tau_1^r+\gamma_1)\\
&= -\mu_1\tau_1^r + \lambda_1\gamma_1\\
&= -\mu_1\left(r\frac{\tau_1^{r-1}}{\mu_1} + r\lambda_1\frac{\lambda_1}{\mu_1}\gamma_1\right) + \lambda_1 r\frac{1}{\lambda_1}\sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}\tau_1^{r-1}}{\mu_2\dots\mu_i}\\
&= -r\tau_1^{r-1} - \lambda_1\frac{r}{\lambda_1}\sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}\tau_1^{r-1}}{\mu_2\dots\mu_i} + r\mu_1\sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}}{\mu_2\dots\mu_i}\\
&= -r\tau_1^{r-1} - \sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}}{\mu_2\dots\mu_i} + \sum_{i=2}^{N}\frac{\lambda_1\dots\lambda_{i-1}}{\mu_2\dots\mu_i}\\
&= -r\tau_j^{r-1}
\end{aligned}$$

Then we continue with the $j-th$ index for $j = 2, \ldots, N-1$:

$$\begin{aligned}
(D\tau^r)_j &= \mu_j \tau_{j-1}^r - (\lambda_j + \mu_j)\tau_j^r + \lambda_j \tau_{j+1}^r \\
&= -\mu_j(\tau_j^r - \tau_{j-1}^r) + \lambda_j(\tau_{j+1}^r - \tau_j^r) \\
&= -\mu_j \frac{r}{\lambda_{j-1}} \sum_{i=j}^{N} \frac{\lambda_{j-1} \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_j \ldots \mu_i} + \lambda_j \frac{r}{\lambda_j} \sum_{i=j+1}^{N} \frac{\lambda_j \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_{j+1} \ldots \mu_i} \\
&= -\mu_j \frac{r}{\lambda_{j-1}} \left(\frac{\lambda_{j-1} \tau_1^{r-1}}{\mu_j} + \sum_{i=j+1}^{N} \frac{\lambda_{j-1} \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_j \ldots \mu_i} \right) + r \sum_{i=j+1}^{N} \frac{\lambda_j \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_{j+1} \ldots \mu_i} \\
&= -r\tau_1^{r-1} - \mu_j \frac{r}{\lambda_{j-1}} \frac{\lambda_{j-1}}{\mu_j} \sum_{i=j+1}^{N} \frac{\lambda_j \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_{j+1} \ldots \mu_i} + r \sum_{i=j+1}^{N} \frac{\lambda_j \ldots \lambda_{i-1} \tau_1^{r-1}}{\mu_{j+1} \ldots \mu_i} \\
&= -r\tau_j^{r-1}
\end{aligned}$$

And finally we show the verification for the $N-th$ index:

$$\begin{aligned}
(D\tau^r)_N &= \mu_N \tau_{N-1}^r - \mu_N \tau_N^r \\
&= -\mu_N(\tau_N^r - \tau_{N-1}^r) \\
&= -\mu_N \gamma_{N-1} \\
&= -\mu_N \frac{r}{\lambda_{N-1}} \sum_{i=N}^{N} \frac{\lambda_{N-1} \ldots \lambda_{i-1} \tau^{r-1}}{\mu_N \ldots \mu_i} \\
&= -\mu_N \frac{r}{\lambda_{N-1}} \frac{\lambda_{N-1} \tau^{r-1}}{\mu_N} \\
&= -r\tau_j^{r-1}
\end{aligned}$$

□

Chapter 4

Numerical examples

4.1 Deterministic model

The deterministic logistic model satisfies the following differential equation

$$\frac{dy}{dt} = \lambda(y) - \mu(y). \tag{4.1}$$

In the logistic model it is reasonable to assume birth and death rates to satisfy

$$\lambda(y) = \lambda_1 y + \lambda_2 y^2 \quad \text{and} \quad \mu(y) = \mu_1 y + \mu_2 y^2$$

where λ_i and μ_i, $i = 1, 2$ are constants such that $\lambda_i \neq \mu_i$. We can write

$$\begin{aligned}
\frac{dy}{dt} &= (\lambda_1 y - \lambda_2 y^2) - (\mu_1 y + \mu_2 y^2), \\
\frac{dy}{dt} &= y(\lambda_1 - \mu_1) - y^2(\mu_2 - \lambda_2), \\
\frac{dy}{dt} &= y(\lambda_1 - \mu_1)\left(1 - y\frac{\mu_2 - \lambda_2}{\lambda_1 - \mu_1}\right), \\
\frac{dy}{dt} &= y(\lambda_1 - \mu_1)\left(1 - \frac{y}{\frac{\lambda_1 - \mu_1}{\mu_2 - \lambda_2}}\right).
\end{aligned}$$

We denote

$$r = \lambda_1 - \mu_1 \quad \text{and} \quad K = \frac{\lambda_1 - \mu_1}{\mu_2 - \lambda_2}$$

where r is the *rate* and K is the *capacity*. Now we can rewrite the equation in the following way

$$\frac{dy}{dt} = yr\left(1 - \frac{y}{K}\right). \tag{4.2}$$

Lemma 4.1.1. *A solution to differential equation (4.1) with initial condition $y(0) = y_0$ is*

$$y(t) = y_0 \frac{Ke^{rt}}{K + y_0(e^{rt} - 1)}, \quad t \geq 0. \tag{4.3}$$

Proof. Proof is given in the Appendix A.2. □

Figure 4.1 and 4.2 show how a deterministic solution can change if we fix r and K and let the initial condition y_0 to change.

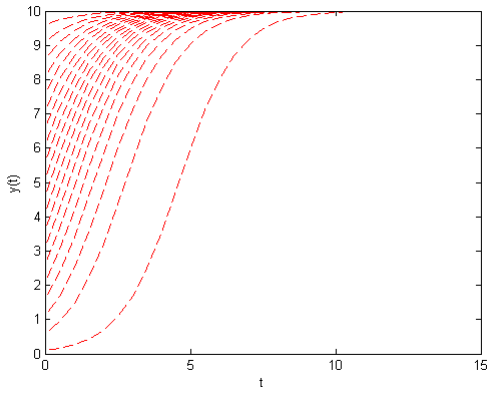

Figure 4.1: Solutions for fixed $r > 0$.

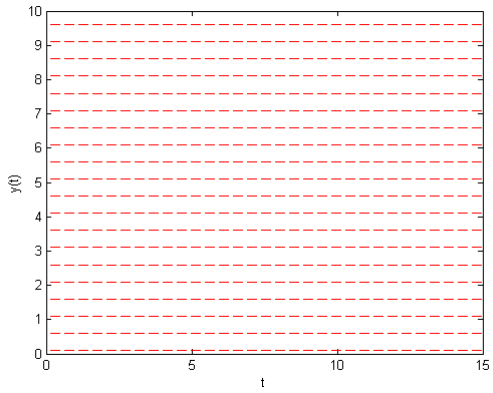

Figure 4.2: Constant solutions ($r = 0$).

Figure 4.1 shows the deterministic solution for $r = 1$ and $K = 10$ and Figure 4.2 for $r = 0$ and $K = 10$. In both cases y_0 changes from 0 to K in steps of $0,5$.

Figure 4.3 and 4.4 also show the deterministic solution for fixed r and K and changing y_0. In 4.4 we take also $y_0 > K$.

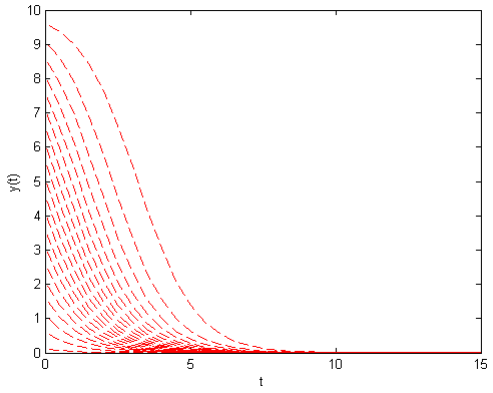

Figure 4.3: Solutions for fixed $r < 0$.

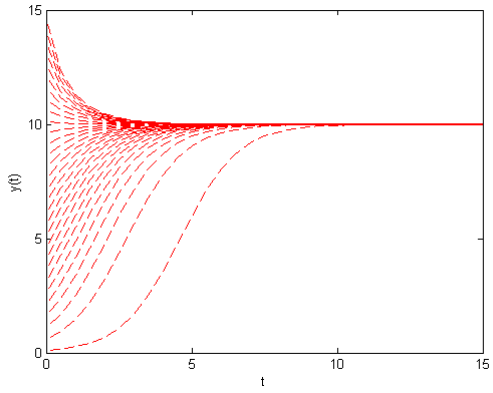

Figure 4.4: Solutions also for $y_0 > K$.

Figure 4.3 shows the deterministic solution for $r = -1$, $K = 10$ and y_0 changing from 0 to K in steps of $0,5$, Figure 4.4 for $r = 1$ and $K = 10$ where y_0 changes from $0,1$ to $K + 5$ in steps of 1.

We show an example where we consider the following birth and death rates. The variable N is the *number of states*.

Example 4.1.1.

$$\lambda(y) = 2y - \frac{y^2}{50} \quad \text{and} \quad \mu(y) = y + \frac{y^2}{50}, \quad N = 100. \tag{4.4}$$

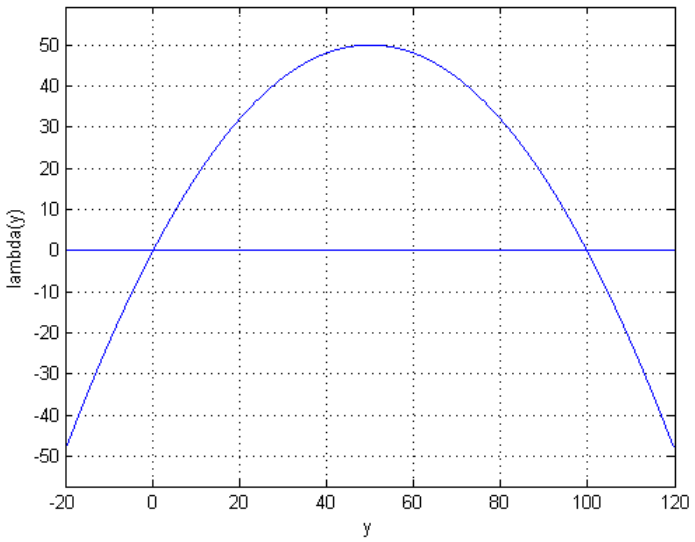

Figure 4.5: Parabola

Birth rate λ is positive only for

$$y(2 - \frac{y}{50}) \geq 0 \Leftrightarrow y \in [0, 100]. \tag{4.5}$$

Natural choice for the maximum number of states (population size) is therefore $N = 100$.

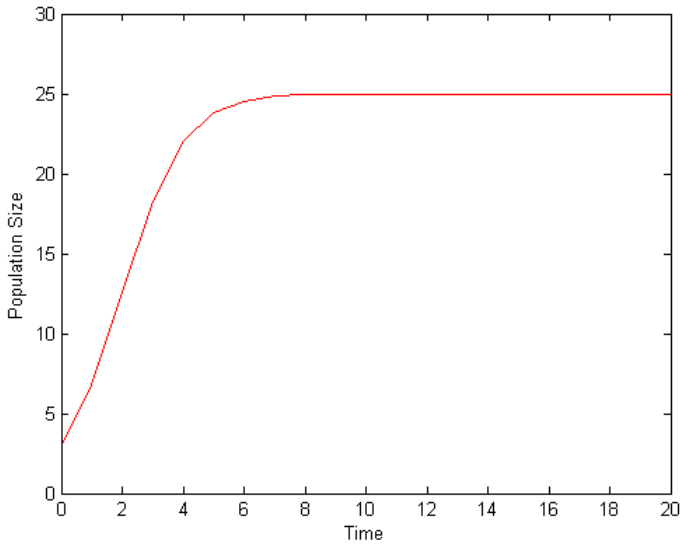

Figure 4.6: Deterministic solution of Example 4.1.1

Figure 4.6 presents the deterministic solution for model with birth and death rates given in (4.4).

4.2 Markov Chain Models

In this section we will compare two Markov chain models, the discrete and the continuous one. We will see that the two models generally agree well even for small population sizes.

Two individual sample paths or stochastic realisations are graphed for each of the two stochastic models and compared to the solution of the deterministic model when the initial population size is $y_0 = 3$.

The two following Figures 4.7 and 4.8 present the sample paths corresponding to the discrete-time Markov chain model and the continuous-time Markov chain model. The blue curves represents individual sample paths, black curve is the mean and two magenta curves represents the mean plus minus the standard deviation. The sample paths are graphed as continuous trajectories although they are not, they are discontinuous at the points where the process jumps to a new state.

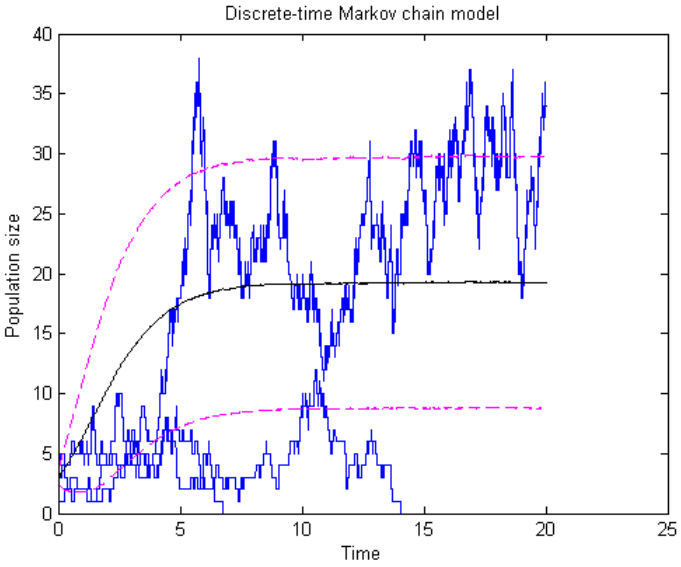

Figure 4.7: DTMC simulation result to Example 4.1.1.

Figure 4.7 shows the discrete-time Markov chain model for Example 4.1.1 with 3 sample paths and the mean of 10000 sample paths. Two trajectories hit zero earlier than $T = 20$, they represent populations that died.

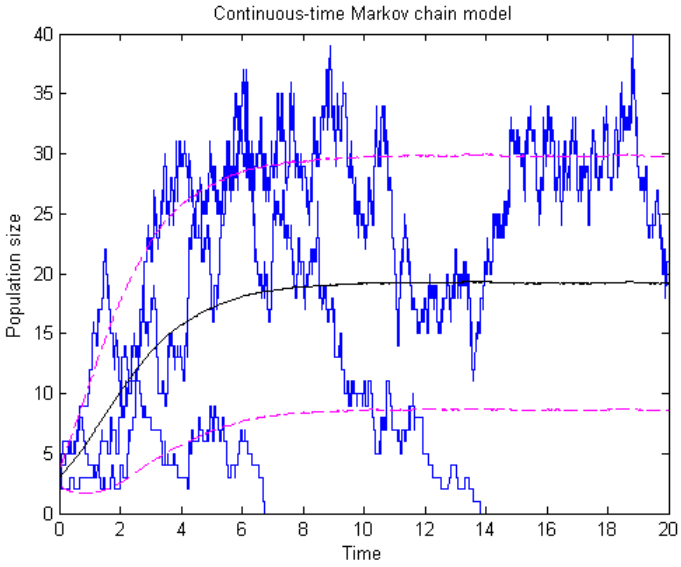

Figure 4.8: CTMC simulation result to Example 4.1.1.

Figure 4.8 shows the continuous-time Markov chain model for Example 4.1.1 with 3 sample paths and the mean of 10000 sample paths. Again two displayed populations died earlier than in $T = 20$.

These two simulations graphed by figures 4.7 and 4.8 show that the dynamics of these two stochastic models are in close agreement even for small initial population sizes.

4.3 Mean Persistence Times

In this section we will compare two models with different birth and death rates. These models will be referred to as Example A and Example B.

$$\lambda(y) = 1.35y - \frac{y^2}{20} \quad \text{and} \quad \mu(y) = 0.35y + \frac{y^2}{20}, \quad N = 27. \tag{A}$$

$$\lambda(y) = 2y - \frac{y^2}{20} \quad \text{and} \quad \mu(y) = y + \frac{y^2}{20}, \quad N = 40. \tag{B}$$

Mean persistence time for Example A and B are calculated using the equation (3.6), equation (3.7) and the simplified equation (3.8) (which give the same output) and compared to simulated results.

The smooth red curve represents the mean persistence time based on calculation mentioned above. The blue points represent the mean persistence times for the discrete-time model and the magenta points represent the mean persistence times for the continuous-time model.

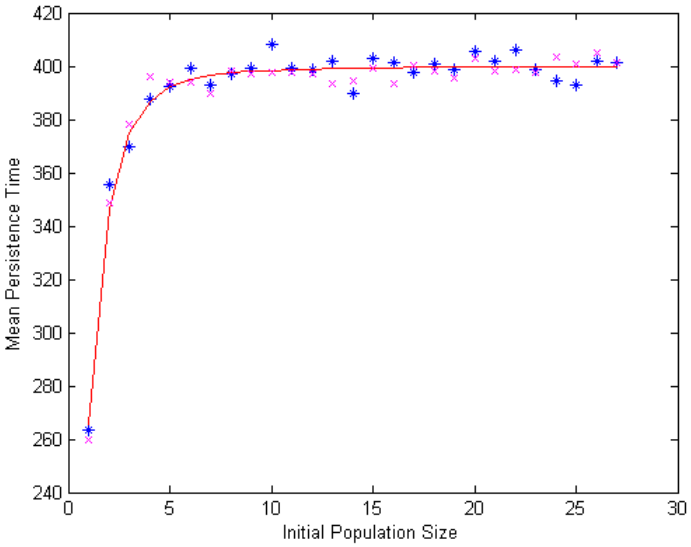

Figure 4.9: Mean persistence times to Example A

Figure 4.9 shows the mean persistence times for Example A both for the discrete-time and the continuous-time model. The individual mean persistence times are computed from 10000 extinct trajectories.

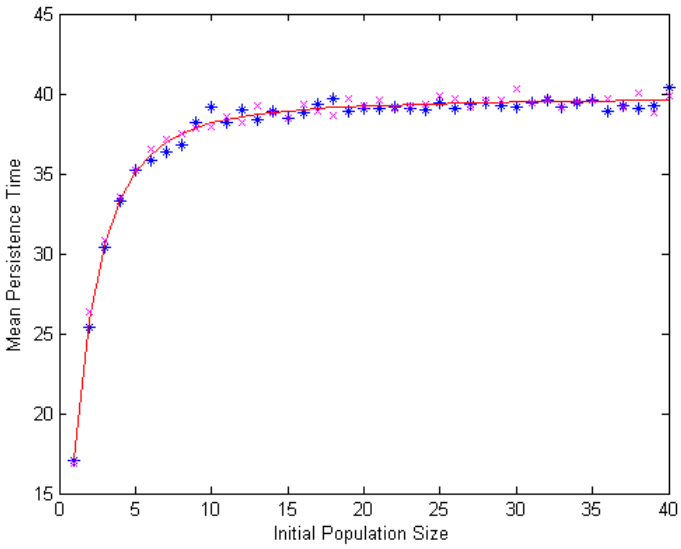

Figure 4.10: Mean persistence times to Example B

Figure 4.10 show the mean persistence time for Example B again both for the discrete-time and the continuous-time model. The individual mean persistence times are computed from 10000 extinct trajectories.

The mean persistence times differs by a factor of 10 in these two examples from a maximum of about 400 in Example A to about 40 in Example B.

The following figures 4.11 and 4.12 compare the two deterministic forms of the solutions of the mean persistence time (3.6) and (3.8). The blue smooth line is the error between the deterministic solutions.

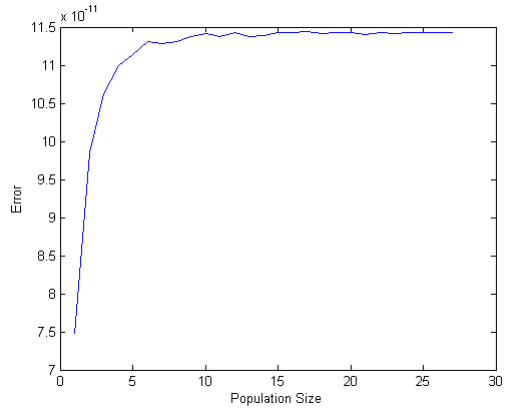

Figure 4.11: Error between (3.6) and (3.8) for Example A

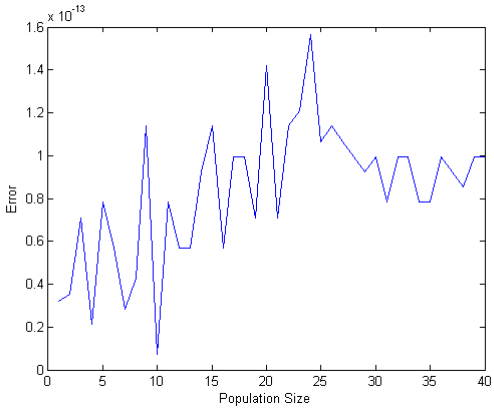

Figure 4.12: Error between (3.6) and (3.8) for Example A

Figure 4.11 is the difference between the deterministic solution for Example A computed using equation (3.6) and (3.8). Figure 4.12 is the difference between the deterministic solution for Example B computed using equation (3.6) and (3.8). The first version for Example A differs at the order of 10^{-11}, the second version for Example B differs at the order of 10^{-13}. These small differences are numerical errors that arose from the matrix inverse in (3.6) or from the recurrent formula (3.8).

In Figures 4.13 and 4.14 the standard deviation for the two stochastic models is graphed. The standard deviation is calculated assuming equation (3.10) for

$$\sqrt{\tau_y^2 - (\tau_y)^2}, \quad r = 2. \tag{4.6}$$

The smooth red curve represents the standard deviation based on calculation mentioned above. The blue curves represent the standard deviations for the discrete-time model and the magenta curves for the continuous-time model. The two green curves are simulated standard deviation $\pm$ deterministic standard deviation computed using equation (4.6).

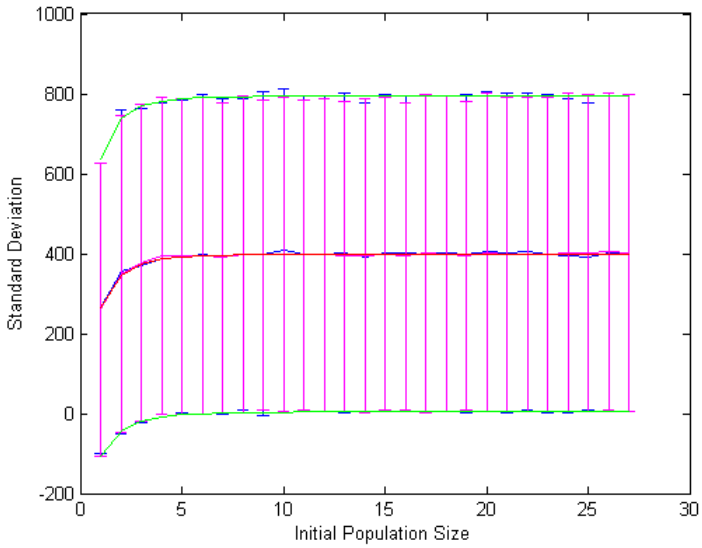

Figure 4.13: Simulation result to Example A

Figure 4.13 shows the standard deviation for Example A both for the discrete-time and the continuous-time model. The individual standard deviations are computed from 10000 extinct trajectories.

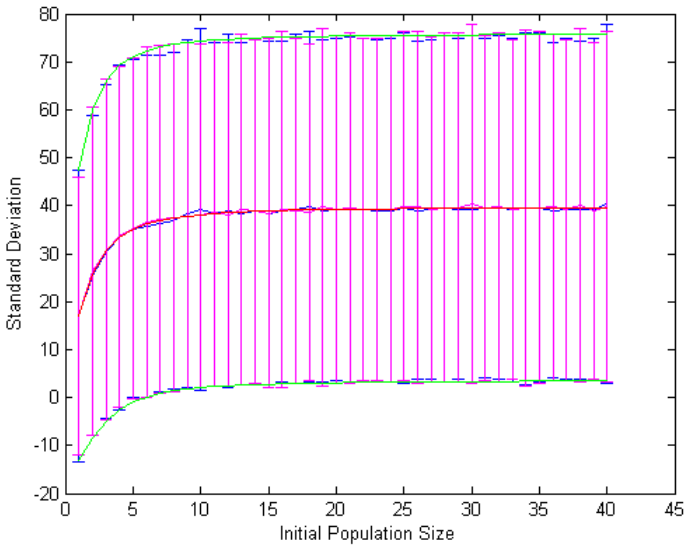

Figure 4.14: Simulation result to Example B

Figure 4.14 shows the standard deviation for Example B again both for the discrete-time and the continuous-time model. The individual standard deviations are computed from 10000 extinct trajectories.

Also here can be seen an order of magnitude difference between the standard deviations in the two cases.

Chapter 5

Conclusion

The majority of stochastic population models in biological applications are continuous in time. However, many population modelling efforts prioritize the discrete-time models, where the populations have non-overlapping generations. In these models, the discrete-time formulations are more biologically realistic.

There exist some advantages of the discrete-time model over the continuous-time model. For example, the discrete-time model is often easier to formulate from the so called first principles (simple discrete rules) and is also simpler understood than the continuous-time models.

Computational ease may be an important consideration in deterministic model behaviour. The discrete-time model has also some numerical advantages over the continuous-time model by projecting forward in time using the transition matrix and calculation of the probability distribution is straightforward. In the continuous-time model the probability distribution can be estimated via simulations of a large number of sample paths. The problem is that to attain a good estimation may require a very large number of sample paths and many more numerical computations than in the discrete-time model. On the other hand, when we consider several interacting populations, the advantages of the discrete-time model may disappear.

The aim of this thesis was to introduce stochastic birth and death processes, summarise the informations obtained from various sources and compare the stochastic processes with the use of simulations. We considered two models, discrete-time and continuous-time Markov chain models. We provided a closed form solution and a recurrent solution for the mean persistence times together with their higher moments, in particular with the variance or standard deviation. We compared the theoretical results with numerical simulations.

Appendix A

Appendix

A.1 The factorisation method for a 3-diagonal matrix

We assume a system of n linear algebraic equations $A * Y = F$ in the following form

$$\left(\begin{array}{ccccc|c} c_1 & -b_1 & & & & f_1 \\ -a_2 & c_2 & -b_2 & & & f_2 \\ & \vdots & \vdots & \vdots & & \\ & & -a_{n-1} & c_{n-1} & -b_{n-1} & f_{n-1} \\ & & & -a_n & c_n & f_n \end{array}\right) \tag{A.1}$$

We denote

$$\alpha_{n-1} = \frac{a_n}{c_n}$$
$$\beta_{n-1} = \frac{f_n}{c_n}.$$

The last two equations can be written in the form (the last one is divided by c_n):

$$-\alpha_{n-1}y_{n-1} + y_n = \beta_{n-1}$$
$$-a_{n-1}y_{n-2} + c_n y_{n-1} - b_{n-1}y_n = f_{n-1}b_{n-1}$$

$$-a_{n-1}y_{n-2} + (c_{n-1} - \alpha_{n-1} - b_{n-1})y_{n-1} = f_{n-1} + \beta_{n-1}b_{n-1}$$

We can rewrite it (the equation divided by the coeficient at y_{n-1}) in the form of

$$\alpha_{n-2}y_{n-2} + y_{n-1} = \beta_{n-2}$$

where

$$\alpha_{n-2} = \frac{a_{n-1}}{c_{n-1} - \alpha_{n-1} - b_{n-1}}$$
$$\beta_{n-2} = \frac{f_{n-1} + \beta_{n-1}b_{n-1}}{c_{n-1} - \alpha_{n-1} - b_{n-1}}.$$

For $n = 3$ (the first equation is multiplied by b_1):

$$\begin{aligned} \alpha_1 y_1 + y_2 &= \beta_1 \\ c_1 y_1 - b_1 y_2 &= f_1 \end{aligned}$$

$$\begin{aligned} y_1(c_1 - \alpha_1 b_1) &= f_1 + \beta_1 b_1 \\ y_1 &= \frac{f_1 + \beta_1 b_1}{c_1 - \alpha_1 b_1} \end{aligned}$$

Then the forward recurrence is:

$$y_n = \beta_{n-1} + \alpha_{n-1} y_{n-1}$$

The algorithm:
1. BACKWARD RECCURENCE:

$$\begin{aligned} \alpha_{n-1} &= \frac{a_n}{c_n} \\ \alpha_i &= \frac{a_{i+1}}{c_{i+1} - \alpha_{i+1} b_{i+1}} \end{aligned}$$

for $i = n-2, n-1, \ldots, 2, 1$.
Where for each right hand side F we calculate:

$$\begin{aligned} \beta_{n-1} &= \frac{f_n}{c_n} \\ \beta_i &= \frac{f_{i+1} + \beta_{i+1} b_{i+1}}{c_{i+1} - \alpha_{i+1} b_{i+1}} \end{aligned}$$

for $i = n-2, n-1, \ldots, 2, 1, 0$.
2. FORWARD RECCURENCE:

$$\begin{aligned} y_1 &= \beta_0 \\ y_i &= \beta_{i-1} + \alpha_{i-1} y_{i-1} \end{aligned}$$

for $i = 2, 3, \ldots, n$.

In our case we assume a system of n linear algebraic equations $A * Y = F$ where A is matrix $-D$ (3.5) and F will be an identity matrix (we take column after column).

In this specific case we know that:

$$\alpha_i = 1$$

for $i = 1, \ldots, n-1$ which very simplifies the calculations.

$$\beta = 0$$

until we get to the row where $f_i = 1$ in the column of the identity matrix. From this point β will be nonzero.

The algorithm:
1. BACKWARD RECCURENCE:

$$\begin{aligned} \alpha_{n-1} &= 1 \\ \alpha_i &= 1 \end{aligned}$$

for $i = n-2, n-1, \ldots, 2, 1$.
Where for each right hand side F we calculate:

$$\begin{aligned} \beta_{n-1} &= \frac{f_n}{\mu_n} \\ \beta_i &= \frac{f_{i+1} + \beta_{i+1}\lambda_{i+1}}{\mu_{i+1}} \end{aligned}$$

for $i = n-2, n-1, \ldots, 2, 1, 0$.
2. FORWARD RECCURENCE:

$$\begin{aligned} y_1 &= \beta_0 \\ y_i &= \beta_{i-1} + y_{i-1} \end{aligned}$$

for $i = 2, 3, \ldots, n$. The factorisation method for a 3-diagonal matrix was derived using the lecture notes for numerical methods [1].

[1] http://trial.zcu.cz/

A.2 The proof of Lemma 4.1.1.

Proof. We check that (4.3) satisfies the differential equation

$$\begin{aligned}
\frac{dy}{dt} &= y_0 \frac{Kre^{rt}[K + y_0(e^{rt} - 1)] - Ke^{rt}[ry_0e^{rt}]}{[K + y_0(e^{rt} - 1)]^2} \\
&= y_0 \frac{K^2re^{rt} - Kry_0e^{rt}}{[K + y_0(e^{rt} - 1)]^2} \\
&= y_0 \frac{(K - y_0)Kre^{rt}}{[K + y_0(e^{rt} - 1)]^2} \\
&= y(t) \frac{(K - y_0)r}{K + y_0(e^{rt} - 1)} \\
&= ry(t) \frac{K + y_0(e^{rt} - 1) - y_0e^{rt}}{K + y_0(e^{rt} - 1)} \\
&= ry(t) \left(1 - \frac{y(t)}{K}\right) \\
y(0) &= y_0 \frac{K}{K + y_0(1 - 1)} = y_0
\end{aligned}$$

Remark A.2.1. To find (4.3) we can use the variation of constants

$$\begin{aligned}
\int \frac{dy}{y\left(1 - \frac{y}{K}\right)} &= r \int dt \\
\int (\frac{1}{y} - \frac{1}{y - K})dy &= r \int dt \\
ln|\frac{y}{y - K}| &= rt + c_1 \\
|\frac{y}{y - K}| &= c_2e^{rt} \\
y &= |y - K|c_2e^{rt} \\
y(c_2e^{rt} - 1) &= Kc_2e^{rt}
\end{aligned}$$

$$y = \frac{Kc_2e^{rt}}{c_2e^{rt} - 1} \tag{A.2}$$

Then we consider the initial condition

$$\begin{aligned}
y_0 &= \frac{Kc_2}{c_2 - 1} \\
y_0(c_2 - 1) &= Kc_2 \\
c_2y_0 - y_0 &= Kc_2
\end{aligned}$$

$$c_2 = \frac{y_0}{y_0 - K} \tag{A.3}$$

Then we set equation (A.3) into (A.2) and we get

$$\begin{aligned} y &= \frac{\frac{Ky_0e^{rt}}{y_0-K}}{\frac{y_0e^{rt}}{y_0-K} - 1} \\ &= \frac{Ky_0e^{rt}}{y_0e^{rt} - y_0 - K} \\ &= y_0\frac{Ke^{rt}}{K + y_0(e^{rt} - 1)} \end{aligned}$$

□

A.3 Description of the thesis attachment

The attached CD contains:

/Matlab/Chapter_/	includes Matlab scripts sorted by their use in chapters.
/Matlab/Data/	mat files with pre–generated data.
/Source_Latex/	LATEX files used for generating the thesis.
/Graphics/Chapter_/	figures that are presented in the thesis sorted by the appearance in chapters (png. formats).

CD root directory:

Latalova_Thesis	the Bachelor thesis.
Readme.txt	this text file.

Bibliography

[1] L.J.S. Allen and E.J. Allen, *A comparison of three different stochastic population models with regard to persistence time*, Theoretical Population Biology (2003), no. 64, 439–449.

[2] P. Bremauld, *Markov chains: Gibbs field, monte carlo simulations and queues*, Springer, New York, 1999.

[3] O. Häggström, *Finite markov chains and algorithmic applications*, Cambridge University Press, New York, 2002.

[4] J.F.C. Kingman, *Markov population processes*, Journal of Applied Probability (1969), no. 6, 1–18.

[5] J.H. Matis and T.R. Kiffe, *Stochastic population models: A comparison of perspective (lecture notes)*, Sheridan Books, New York, 2000.

[6] Z. Prášková and Lachout P., *Základy náhodných procesů*, Karolinum, Praha, 2001.

[7] M. Uyenoyama and R. Singh, *The evolution of population biology*, Cambridge University Press, New York, 2004.

[8] D. Worster, *Nature´s economy*, Cambridge University Press, New York, 1994.